INL/CON-07-12772
PREPRINT

Meeting Human Reliability Requirements Through Human Factors Design, Testing, and Modeling

Proceedings of the European Safety and Reliability Conference (ESREL 2007)

R. L. Boring

June 2007

The INL is a
U.S. Department of Energy
National Laboratory
operated by
Battelle Energy Alliance

Meeting Human Reliability Requirements through Human Factors Design, Testing, and Modeling

R.L. Boring
Idaho National Laboratory, Idaho Falls, Idaho, USA
and
OECD Halden Reactor Project, Halden, Norway

ABSTRACT: In the design of novel systems, it is important for the human factors engineer to work in parallel with the human reliability analyst to arrive at the safest achievable design that meets design team safety goals and certification or regulatory requirements. This paper introduces the System Development Safety Triptych, a checklist of considerations for the interplay of human factors and human reliability through design, testing, and modeling in product development. This paper also explores three phases of safe system development, corresponding to the conception, design, and implementation of a system.

1 INTRODUCTION

Superficially, human factors and ergonomics (HFE) and human reliability analysis (HRA) appear to be—if not identical—at least complementary fields. HFE combines interdisciplinary elements of engineering, psychology, and computer science, among other fields, into a cohesive discipline (Boring, 2002). Likewise, HRA features the cornerstones of engineering and psychology, coupled with elements of risk assessment. Despite these surface similarities, the two fields have markedly different core areas of focus. Specifically, whereas HFE has emerged as an integral part of design and engineering disciplines, HRA is an integral part of probabilistic safety assessment (PSA). Therein resides the key difference between HFE and HRA. HFE is heavily involved in the design of novel systems, whether for usability (Nielsen, 1993), enjoyment (Norman, 2002), or safety (Palanque et al., 2004). HRA is focused on verifying the safe performance of human actions, often in the context of the retrospective analysis of incidents, events, or accidents (Gertman and Blackman, 1994).

The line between HFE and HRA is blurred when considering the design of novel safety-critical systems that must be human certified. This possibility becomes relevant as the next generation of nuclear power plant control rooms (Boring et al., 2005), aerospace systems (NASA, 2005), and air traffic control systems (Abbott et al., 2006) are designed and deployed. Each of these fields sets a high bar for safety—safety that meets or exceeds current technology; at the same time, each of these fields demands innovative human-machine interfaces that enable the human to accomplish more in a simpler fashion.

In some cases, the integration of HRA into HFE is regulated. For example, NASA Procedural Requirement (NPR) 8705.2A, *Human-Rating Requirements for Space Systems* (2005), identifies the process by which a hardware or software system for space use may become Human-Rating certified. The Human-Rating Requirements apply to any hardware or software space system that is developed and/or operated by or for NASA and that supports humans or interacts with another system that supports humans. The purpose of the Human-Rating Requirements is to ensure that no single system (i.e., single-point) failure and no two inadvertent actions cause death or permanent disability to public, crew, passengers, or ground personnel.

The present document provides guidance for HFE practitioners to work alongside risk analysts to meet safety goals and regulatory requirements in the development of novel systems. By design, this document does not rigidly specify the acceptance criteria for a system to be considered safe or certified. Rather, this document specifies a process that may be followed to ensure the best achievable safety for a novel system.

2 BASIC SAFETY DESIGN PROCESS

Development of a human-system interface (HSI) that meets safety and human reliability goals involves a phased process from conception to implementation. Safety is demonstrated through three successive phases that should be treated with independent milestones and objectives. Phase I, which

consists of the conceptual or specification phase of system development, should outline not only desirable system features but also planned compliance to applicable safety standards throughout the lifecycle of the system. This phase may include setting safety objectives in terms of HRA such as, for example, the maximum acceptable human error probability of 1E-3 for any single action. Phase II, which is the preliminary design or prototype phase, is most closely associated with HFE, in that initial designs are prototyped and tested prior to product implementation. In this phase, HRA may be used to produce evidence that the safety objectives have been met. During Phase II, the HFE engineer also identifies and reviews critical functions for system operation and personnel safety. Phase III corresponds to the actual implementation of the HSI. During the phase, the HFE engineer compiles a description of the critical function performance criteria through analysis, testing, inspection, and documentation.

3 ACHIEVING CERTIFICATION

As noted previously, some HRA-infused HFE serves specific regulatory requirements. At what point has the HSI reached the acceptable, safe level of human reliability? Regulatory approval or certification may occur at any or all of the three phases. The guidance documents overviewed in the certification process (e.g., the NASA *Human-Rating Requirements*) typically provide proven methods to model, design, test, and validate safety-critical HSI. Because of the diversity of systems and applications, it is also beyond the scope of the requirements to specify the point at which any given system may be certified. It is for this reason that a safety process, as prescribed here, can help ensure safety considerations have been adequately addressed in the HFE process.

In some cases, where safety considerations are not practicable, waivers or deviations may be granted. These waivers or deviations may only be granted if best achievable safety levels have been met and there is a strong basis for the waiver or deviation such as technical infeasibility or reasonable cost limitations of the system. Importantly, a waiver or deviation does not obviate the need to meet best achievable safety levels. A waiver or deviation, in fact, certifies that every feasible and reasonable step has been taken to ensure the safety of those humans who come in contact with the HSI.

Neither infusing HRA into the HFE process nor achieving certification guarantees that a system is actually safe in all cases. A carefully designed and reviewed system may fail to consider all possible scenarios that could contribute to human error or unsafe interaction with a system. The literature is rife with examples of incidents, events, and accidents that resulted from otherwise well designed systems but that failed to consider the full spectrum of possible safety-degrading scenarios with which the users of those systems would be confronted (Johnson, 2003). The inability to predict such scenarios is at the heart of resilience engineering (Hollnagel, 2006). Resilience in this sense embarks upon the task of anticipating risks, even in the face of unpredictable and dynamic circumstances. One key to this approach is to avoid using predefined sets of risks (e.g., risk cutsets as used in PSAs) in designing and certifying a safe system. Rather, the goal of a resilient safety process is to understand the human response in the face of risks. The process proposed in this paper is congruent with resilience engineering in that it proposes an integrative and iterative merger of design, testing, and modeling. This approach can be used to address specific pre-defined risks according to regulatory guidelines. It can, however, also be used to develop a conceptual understanding of how humans using a particular system respond to off-normal or suboptimal scenarios. Such insights extend beyond addressing pre-defined risks; rather, they provide a catalog of human actions and inactions—both safe and unsafe—and a basis for mitigating unsafe actions and enhancing safe actions.

4 BEST ACHIEVABLE PRACTICES

In this section, I outline steps that may be used to guide the best achievable practices in safety-critical HFE. These steps may also help achieve safety goals and acceptance criteria for certification. The factors that determine the best achievable safety practices include HFE design, testing, and modeling, and together comprise the System Development Safety Triptych. Note that design and testing are largely the domain of traditional HFE, while modeling adds HRA considerations. It is the union of human factors design and testing with human reliability modeling that can help realize best achievable safety in the product development process. Note that these considerations are not prescriptive of a normative model of design for safety. They serve merely as suggestions gleaned from the author's observations in integrating HFE and HRA.

4.1 *Best Achievable Practices for Design*

The following practices, including hardware and software engineering, may facilitate safety-critical design.

- *Compliance with applicable standards and best practices documents.* Applicable standards will vary according to the specific system that is being designed. Where applicable, ANSI, ASME, IEEE, ISO, or other discipline-specific standards and best practices should be followed carefully.

Table 1. The System Development Safety Triptych of design, testing, and modeling.

Design *(Including hardware and software engineering)*	**Testing** *(including equipment and human subject testing)*	**Modeling** *(including PSA, HRA, and simulations)*
• Compliance with applicable standards and best practices documents	• Controlled studies that avoid confounds or experimental artifacts	• Compliance with applicable standards and best practices documents
• Consideration of system usability and human factors	• Use of maximally realistic and representative scenarios, users, and/or conditions	• Use of established modeling techniques
• Iterative design-test-redesign-retest cycle	• Use of humans-in-the-loop testing	• Validation of models to available operational data
• Tractability of design decisions	• Use of valid metrics such as statistically significant results for acceptance criteria	• Completeness of modeling scenarios at the correct level of granularity
• Verified reliability of design solutions	• Documented test design, hypothesis, manipulations, metrics, and acceptance criteria	• Realistic model end states

Where deviation is necessary due to the unique nature of particular systems, these deviations should be clearly documented and justified.

- *Consideration of system usability and human factors.* If the system that is being designed will be used by a human, it should be designed according to usability and human factors standards such as NASA-STD-3000, MIL-STD-1472, and ISO 9241. The system should be ergonomically designed to facilitate the user's natural physical interaction with the system. Where applicable, this consideration includes unique environmental considerations such as human-system interaction in a weightless environment. The system should be designed to be maximally usable in terms of the quality of the software or hardware interface. Best practices regarding the use of display, interface, and control design should be followed and documented. In terms of safety-critical certification, the primary emphasis of usability and human factors is on ensuring that the use of a system does not disrupt critical functions or compromise the safety of the user or other humans.

  Systems that are not directly used by humans but that may interact with human-occupied systems (HOS) should ensure that the non HOS will not change or disrupt the normal human-system interaction of the HOS.

- *Iterative design-test-redesign-retest cycle.* The complement of good design is testing the first-effort design and applying lessons learned in the refinement of the design. Where feasible, system design should be tested to identify potential issues in terms of critical functions. Compliance with design standards does not guarantee a system optimized for human safety. Iterative testing and redesign of a system throughout the design lifecycle helps demonstrate safety in the design.

- *Tractability of design decisions.* Where decisions have been made that could affect the critical functions of the system, these decisions should be clearly documented. Design decisions that could adversely impact critical functions under any circumstances must be justified. As well, all safety considerations should be documented (e.g., the system component may fail under unusual circumstances; however, a different system component ensures the viability of the personnel while repairs can be made). This design rationale may serve as the basis for waivers or deviations from certification.

  Documentation of design decisions that enhance critical functions serves to facilitate the safety goals and certification process.

- *Verified reliability of design solutions.* The reliability of systems should be documented through vendor data, cross-reference to the operational

history of similar existing systems, and/or test results. A system's reliability may be modeled through a composite of sub-component reliability data, but whole-system testing is generally preferable. Acceptable reliability levels should be identified and agreed upon early in the design process. It is especially important to project system reliability throughout the system lifecycle, including considerations for maintenance once the system has been deployed. It is also important to incorporate the estimated mean time before failure into the estimated life of the system.

4.2 *Best Achievable Practices for Testing*

The following testing practices, including equipment and human subject testing, may facilitate the overall safety of the system.

- *Controlled studies that avoid confounds or experimental artifacts.* Testing of system designs should be accomplished using rigorous experimental methods specific to the application and system. Testing may include hardware reliability testing, HSI usability evaluation, and software debugging. Testing should avoid situations that could lead to ambiguous results, such as when alternate causal explanations (e.g., confounds) or an unrealistic experimental design (e.g., experimental artifacts) come into play.

- *Use of maximally realistic and representative scenarios, users, and/or conditions.* To the extent feasible, testing should involve a real-time, closed-loop test environment. The testing scenarios and conditions should reflect the range of actions the system will experience in actual use, including possible worst-case situations. Similarly, human subject testing should involve users who are representative of the actual system users under environmental factors and situations characteristic of the expected usage.

- *Use of humans-in-the-loop testing.* A system that will be used by humans should always be tested by humans. Per the previous point, the human subjects should be representative of actual users. All human subject testing should carefully follow and document safety and ethics standards for treatment of human participants. Hazardous environments should be simulated to the extent feasible and safe for testing purposes.

- *Use of valid metrics such as statistically significant results for acceptance criteria.* Testing should be measured using methods appropriate to the discipline. Where feasible, the metrics should reflect system or user performance across the entire range of expected circumstances. In many cases, testing will involve use of a statistical sample evaluated against a pre-defined acceptance (e.g., alpha) level for "passing" the test. Inferential statistics are preferable to descriptive statistics. Inferential statistical analyses should clearly define the acceptance level (e.g., $\alpha \leq 0.001$). Care should be taken to ensure proper statistical power and to avoid Type I errors (i.e., false positives).

- *Documented test design, hypothesis, manipulations, metrics, and acceptance criteria.* A well documented test plan is crucial in the initial phases of the safety design process. This test plan should include the test design, hypothesis (or hypotheses), manipulations, metrics, and acceptance criteria. Subsequent testing should closely follow this test plan and document any required deviations from the test plan. In the event that a system fails to meet the identified acceptance criteria during testing, a redesign and additional testing should be undertaken before proceeding with final implementation.

4.3 *Best Achievable Practices for Modeling*

The following modeling practices, including PSA, HRA, and simulations, may facilitate the overall safety of the system.

- *Compliance with applicable standards and best practices documents.* Regulatory and standards agencies provide guidance across the domains of PSA—e.g., NASA NPR 8705.5, *Probabilistic Risk Assessment (PRA) Procedures for NASA Programs and Projects* (2004)—and HRA—e.g., US Nuclear Regulatory Commission NUREG-1792, *Good Practices for Implementing Human Reliability Analysis* (2005). Specific modeling techniques with established histories will typically have documentation to outline best practices. Where deviation is necessary due to the unique nature of particular systems, these deviations should be clearly documented and justified.

- Use of established modeling techniques. Within PSA, HRA, and simulation modeling, many established and novel techniques exist. For certification purposes, it is better to use an existing, vetted method than to make use of novel techniques and methods that have not been established within a particular industry. Modeling tools and techniques should be documented early in the safety design process to ensure concurrence by review agencies.

- *Validation of models to available operational data.* To ensure a realistic modeling representation, models must be baselined to data obtained

from empirical testing or actual operational data. Such validation increases the veracity of model extrapolations to novel domains. For example, if a system component has been tested in a particular environment and a simulation model accurately reflects performance in that environment, it may be acceptable to extrapolate the model to a novel environment in which the component has never been tested but in which the performance parameters are predicted to follow a well understood pattern. Note that it is generally preferable to conduct an actual test on novel performance situations rather than to model that performance.

- *Completeness of modeling scenarios at the correct level of granularity.* Modeling scenarios should cover the complete range of operating scenarios with reasonable concession for the possibility of negative outcomes. A thorough task analysis, a review of relevant past operating experience, and a review by subject matter experts help to ensure the completeness of the model.

  The appropriate level of task decomposition or granularity should be determined according to the modeling method's requirement, the fidelity required to model success and failure outcomes, specific requirements of the system that is being designed, and guidance provided by regulatory boards, as appropriate. Every step or action that can affect system success and personnel safety should be modeled. Where steps or actions are logically related, they may be clustered as a single step or action, provided this grouping does not mask opportunities for failure or recovery steps.

- *Realistic model end states.* A PSA or HRA model or simulation must provide a realistic set of situations to which it is modeled. The end states should reflect reasonable and realistic outcomes across the range of operating scenarios. These end states should reflect possible negative outcomes such as system failure. Negative outcomes should always ensure the opportunity for reasonable recovery of humans interacting with that system.

5 BEST ACHIEVABLE PRACTICES ACROSS SYSTEM DEVELOPMENT PHASES

Table 2 outlines how the best achievable practices for design, testing, and modeling may be implemented across the safety design process. In Phase I, the planned use or implementation of best achievable practices is documented. In Phase II, the actual use or implementation of these practices is documented. Finally, in Phase III, the results or products across design, testing, and modeling are documented. At each phase of the safety design process, the HFE design team should carefully review the progress in meeting the dual goals of best achievable safety practices and usable product design.

At the end of the safety design process, the HFE design team and/or a regulatory agency may review the quality of the overall design in terms of safety. The responsible parties should determine one of three possible outcomes for the system:

- *Unsafe.* If the system has failed to meet best achievable practices and the specific safety requirements set forth by the team or the regulator, the system should *not* be considered safe without redesign efforts.

- *Waiver or Deviation.* If the system has met best achievable practices but cannot reasonably or realistically meet the specific safety requirements set forth by the team or the regulator, the system may receive a waiver or deviation from the original safety plan or the certification process.

- *Safe.* If the system has met best achievable practices and the specific safety requirements set forth by the team or the regulator, the system may be considered safe within the specified safety parameters.

6 DISCUSSION

The System Development Safety Triptych process outlined in this document is illustrative of the value of fusing HFE with HRA. Traditional HFE involvement in the development process encompasses design and testing. While these phases help ensure safe systems, they do not individually or collectively meet all the considerations of a safe system. When augmented with modeling considerations of HRA and PSA, however, they can more effectively meet safety goals and regulatory and certification requirements. Similarly, HRA and PSA, when omitting the design and testing contributions of HFE, can fail to provide the input essential to designing and validating a safe system. It is the interplay of both HFE and HRA that most effectively leads to safe, even certifiably safe, systems.

7 DISCLAIMER

This paper represents the author's interpretation of a process that may address safety and design requirements. The author is not affiliated with US or inter-

Table 2. Best achievable practices for safety across development phases.

	Best Achievable Practice	Phase I Conceptual or Functional Specification	Phase II System Design	Phase III System Implementation	Review Outcomes
Design	*Compliance with applicable standards and best practices documents*	Planned Use	Actual Use	Product/Results	
	Consideration of system usability and human factors	Planned Use	Actual Use	Product/Results	
	Iterative design-test-redesign-retest cycle	Planned Use	Actual Use	Product/Results	**Unsafe** - Have not met best achievable safety practices - Have not met all safety goals or regulatory requirements
	Tractability of design decisions	Planned Use	Actual Use	Product/Results	
	Verified reliability of design solutions	Planned Use	Actual Use	Product/Results	
Testing	*Controlled studies that avoid confounds or experimental artifacts*	Planned Use	Actual Use	Product/Results	
	Use of maximally realistic and representative scenarios, users, and/or conditions	Planned Use	Actual Use	Product/Results	**Waiver or Deviation** - Have met best achievable safety practices - Have not met all safety goals or regulatory requirements
	Use of humans-in-the-loop testing	Planned Use	Actual Use	Product/Results	
	Use of valid metrics such as statistically significant results for acceptance criteria	Planned Use	Actual Use	Product/Results	
	Documented test design, hypothesis, manipulations, metrics, and acceptance criteria	Planned Use	Actual Use	Product/Results	
Modeling	*Compliance with applicable standards and best practices documents*	Planned Use	Actual Use	Product/Results	**Safe** - Have met best achievable safety practices - Have met all safety goals or regulatory requirements
	Use of established modeling techniques	Planned Use	Actual Use	Product/Results	
	Validation of models to available operational data	Planned Use	Actual Use	Product/Results	
	Completeness of modeling scenarios at the correct level of granularity	Planned Use	Actual Use	Product/Results	
	Realistic model end states	Planned Use	Actual Use	Product/Results	

national regulatory agencies and has neither implied nor explicit endorsement by such agencies for thework presented in this paper. This research was carried out by Idaho National Laboratory, a US Department of Energy laboratory operated by Battelle Energy Alliance. This report was prepared as an account of work sponsored by an agency of the US Government. Neither the US Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product, or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

8 REFERENCES

Abbott, K., Corker, K., Figarol, S., Hoekstra, J., Prichett, A., & Sheridan, T. 2006. Panel on human-automation interaction in the next generation air transportation system. *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero 2006)*, 246-247.

Boring, R.L. 2002. Human-computer interaction as cognitive science. *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*, 1767-1771.

Boring, R., Hugo, J., Richard, C., & Dudenhoeffer, D. 2005. The role of human-computer interaction in next-generation control rooms. *CHI 2005 Conference Companion*, 2033-2034.

Gertman, D.I., & Blackman, H.S. 1994. *Human Reliability & Safety Analysis Data Handbook.* New York: Wiley-Interscience.

Hollnagel, E. 2006. Resilience—the chanllenge of the unstable. In E. Hollnagel, D.D. Woods, & N. Leveson (eds.), *Resilience Engineering: Concepts and Precepts.* Burlington, VT: Ashgate Publishing Company.

Johnson, C.W. 2003. *Failure in Safety-Critical Systems: A Handbook of Accident and Incident reporting.* Glasgow: University of Glasgow Press.

National Aeronautical and Space Administration. 2004. *Probabilistic Risk Assessment (PRA) Procedures for NASA Programs and Projects, NPR 8705.5.* Washington, DC: NASA Office of Safety and Mission Assurance.

National Aeronautical and Space Administration. 2005. *Human-Rating Requirements for Space Systems, NPR 8705.2A.* Washington, DC: NASA Office of Safety and Mission Assurance.

Nielsen, J. 1994. *Usability Engineering.* San Francisco: Morgan Kaufman.

Norman, D.A. 2002. Emotion and design: Attractive things work better. *Interactions Magazine, IX* (4): 36-42.

Palanque, P., Johnson, C., Koornneef, F., Szwillus, G., & Wright, P. 2004. Safety-critical interaction: Usability in incidents and accidents. *CHI 2004 Extended Abstracts*, 1600-1601.

US Nuclear Regulatory Commission. 2005. *Good Practices for Implementing Human Reliability Analysis (HRA), NUREG-1792.* Washington, DC: US Nuclear Regulatory Commission.